Using EEG-validated Music Emotion Recognition Techniques to Classify Multi-Genre Popular Music for Therapeutic Purposes

Dejoy Shastikk Kumaran, NUS High School

Acknowledgements:

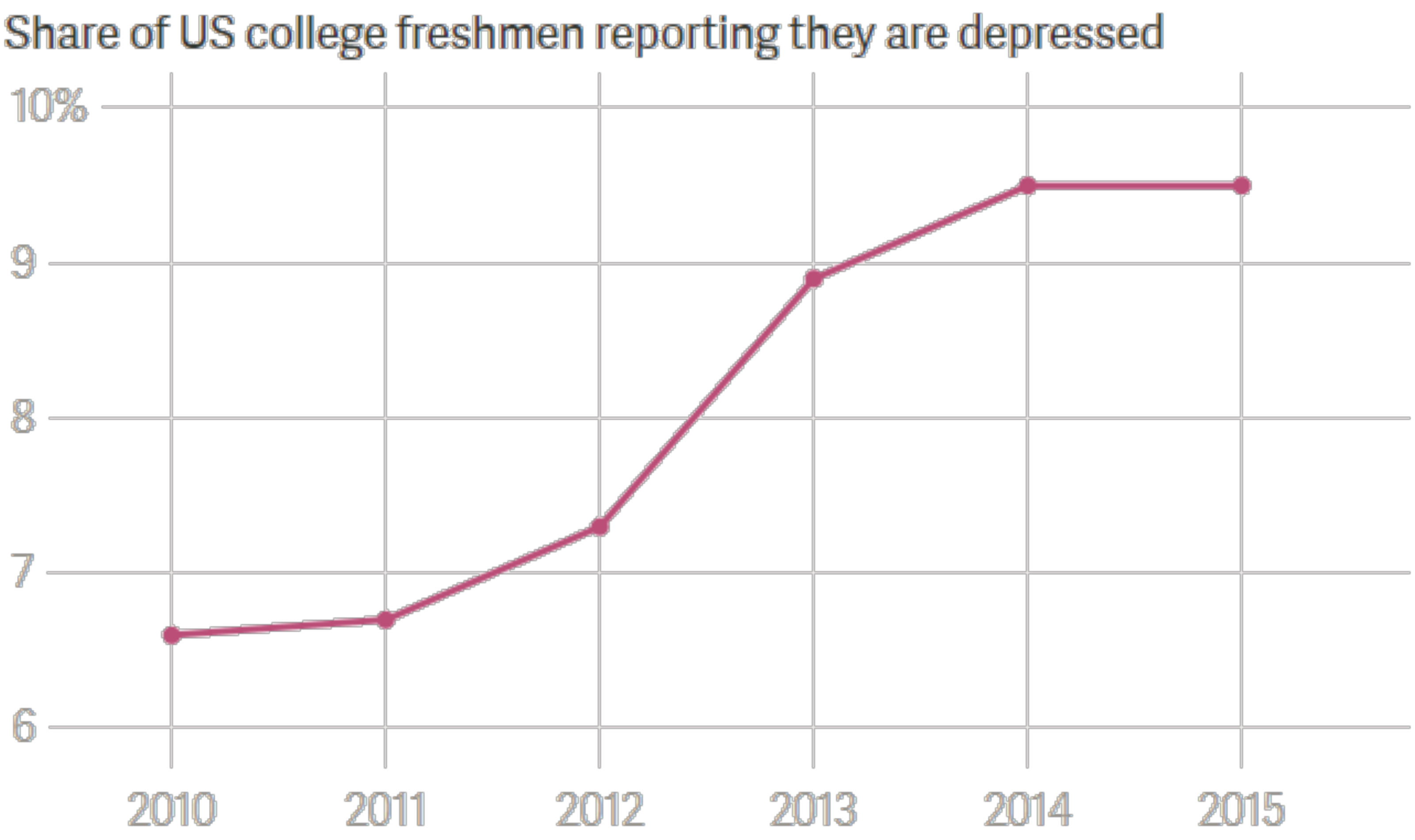
Saravanan Yukesh Ragavendar, NUS High School
Aung Aung Phyo Wai, Nanyang Technological University
Liu Xiaoyu, Apple Inc.
Mr. Claude Chua, NUS High School

1

INTRODUCTION

"…lowering apical heart rates and raising peripheral temperatures were more successful in the relaxation and music therapy groups than in the control group. The incidence of cardiac complications was found to be lower in the intervention groups…"

(Guzetta, 1989)



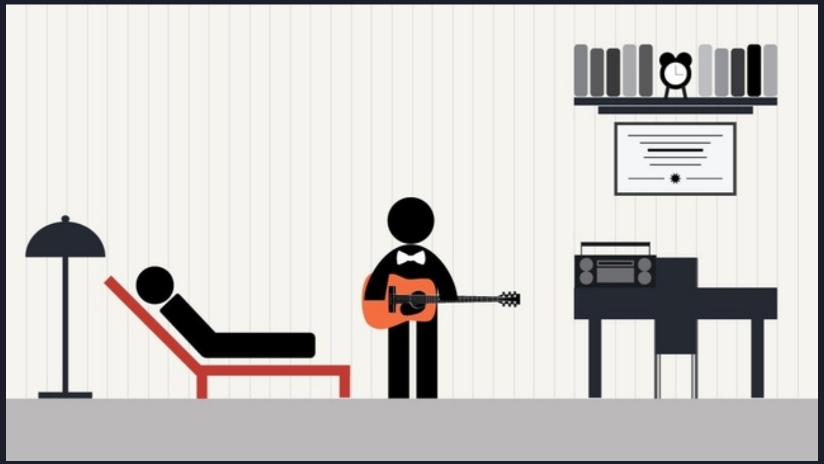
Increasing numbers of U.S. college freshmen and those in high school are reporting depression and high levels of stress**.**

This has been coupled with increasing abuse of Xanax and other Benzos, which has caused a **teenage health crisis** amongst teenagers.

Music Therapy has potential to reduce depression and stress, but…

Subjective Response

Objective Music Features
(Generalizable!)



Most existing studies **have a small sample size** and **subjective response** experimental techniques and thus any specific findings on the **implementation** of music therapeutic programs are not broadly generalizable.

Is there a scientific way to use the features extracted from audio files to predict music emotion?

Machine Recognition of Music Emotion: A Review

YI-HSUAN YANG and HOMER H. CHEN, National Taiwan University

The proliferation of MP3 players and the exploding amount of digital music content call for novel ways of music organization and retrieval to meet the ever-increasing demand for easy and effective information access. As almost every music piece is created to convey emotion, music organization and retrieval by emotion is a reasonable way of accessing music information. A good deal of effort has been made in the music information retrieval community to train a machine to automatically recognize the emotion of a music signal. A central issue of machine recognition of music emotion is the conceptualization of emotion and the associated emotion taxonomy. Different viewpoints on this issue have led to the proposal of different ways of emotion annotation, model training, and result visualization. This article provides a comprehensive review of the methods that have been proposed for music emotion recognition. Moreover, as music emotion recognition is still in its infancy, there are many open issues. We review the solutions that have been proposed to address these issues and conclude with suggestions for further research.

Categories and Subject Descriptors: H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval—*Retrieval models*; H.5.5 [**Information Interfaces and Presentation**]: Sound and Music Computing—*Systems, modeling*

General Terms: Algorithms, Performance, Human Factors

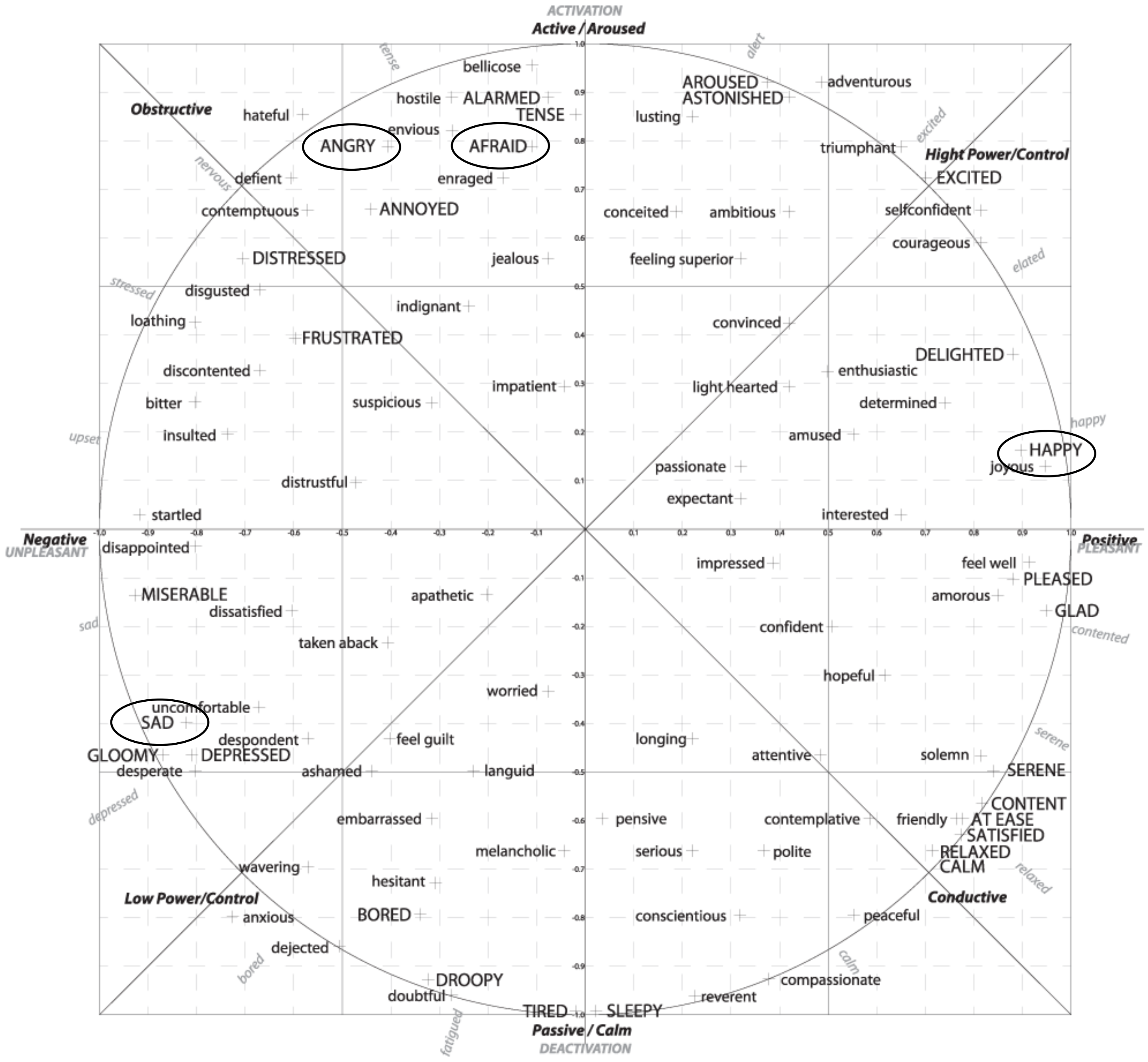
Additional Key Words and Phrases: Music emotion recognition

Machine Recognition of Music Emotion is a technique that uses **classification** of **extracted audio features** (or language features, if NLP is used) to classify music into **various emotions.**

1. Machine Recognition of Music Emotion has **low accuracy rates** in existing literature, with accuracies below 66% for studies using **only low-level features**

2. Machine Recognition of Music Emotion has not been applied to **popular music**, and has mostly been applied to obscure music which is of **less public relevance.**

3. Machine Recognition of Music Emotion has not been applied to **popular music**, and has mostly been applied to obscure music which is of **less public relevance.**



Valence – How positive or negative an emotion is

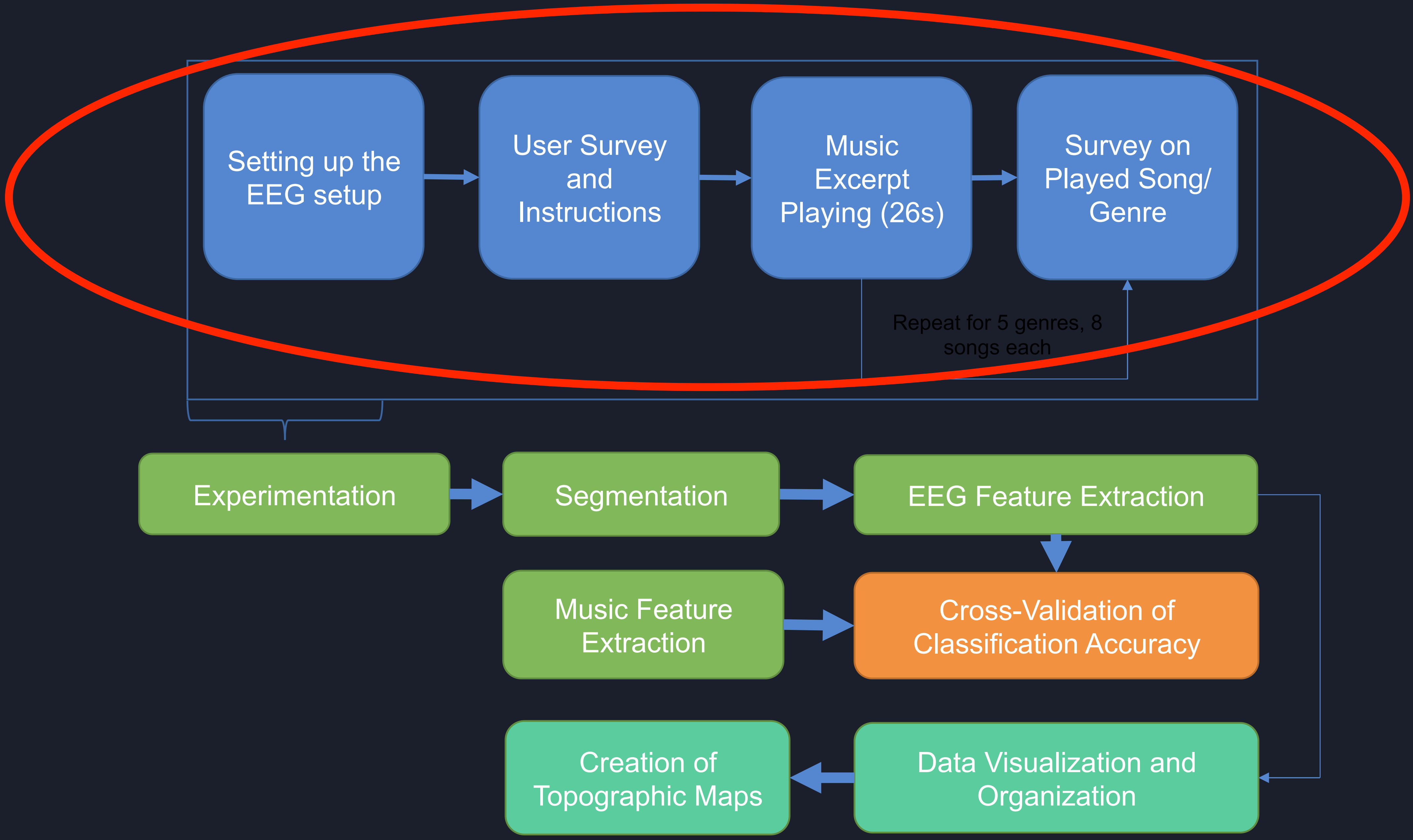
Arousal – How intense an emotion is

We are only studying the circled emotions – (happy, angry, afraid and sad) in this study. (Ask me why later!)

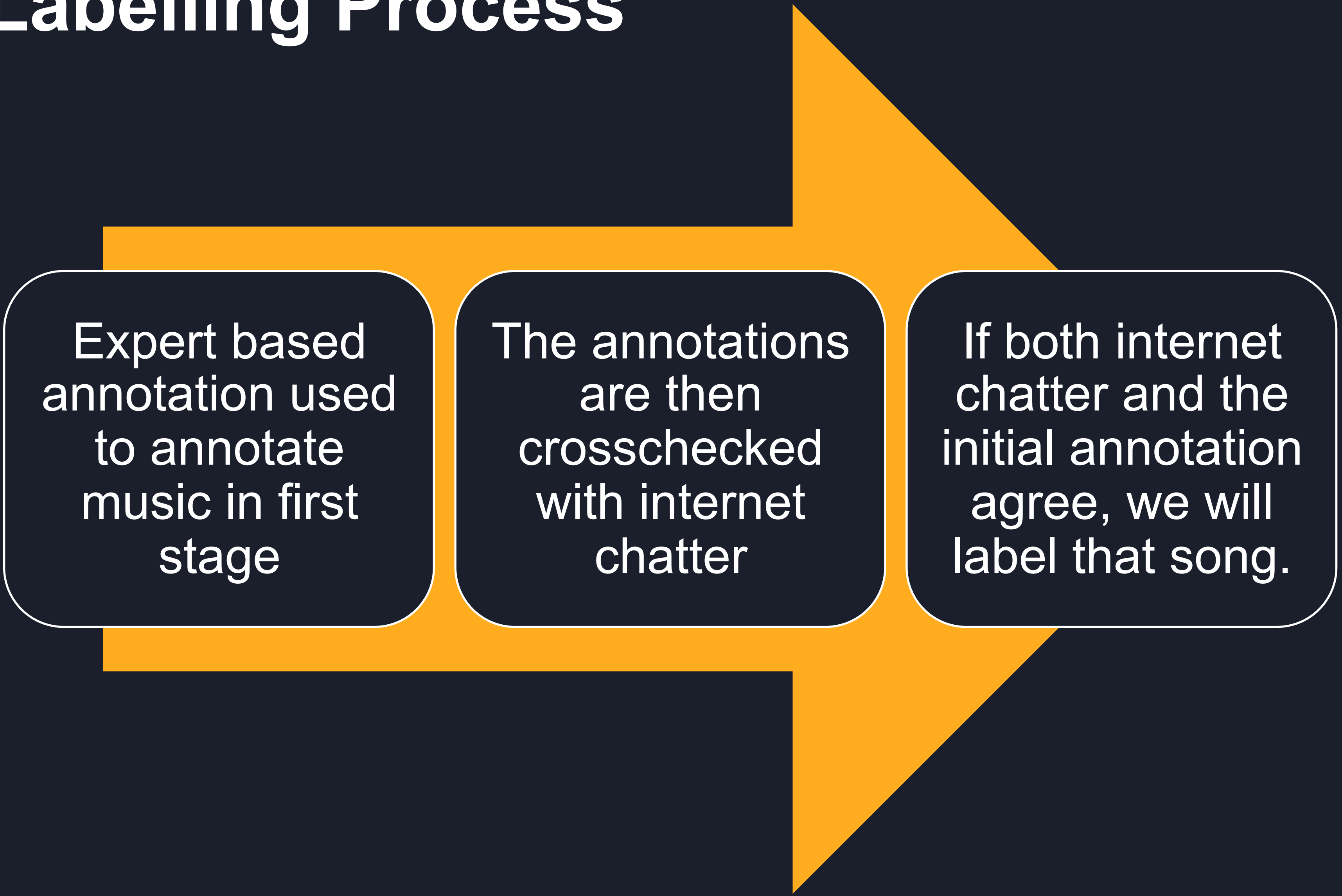
1. Obtain **subject-independent** classification accuracies for **EEG** and **Music Feature** data
2. **Experimentally validate** the viability of **subjective annotation** in determining music emotion
3. **Experimentally validate** the viability of **machine recognition** of music emotion
4. Analysing **subjective annotations** and **topographic maps** for further use
5. Design methods for **emotional induction** and classify **2500** songs by emotion
6. Implement **validated application** allowing for **genre specific** playlist creation

2

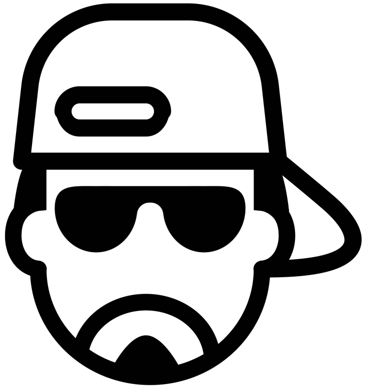
EXPERIMENTATION


Setting up the EEG setup
User Survey and Instructions
Music Excerpt Playing (26s)
Survey on Played Song/ Genre
Repeat for 5 genres, 8 songs each
Experimentation
Segmentation
EEG Feature Extraction
Music Feature Extraction
Cross-Validation of Classification Accuracy
Creation of Topographic Maps
Data Visualization and Organization

Emotion Labelling Process



Selection of Songs



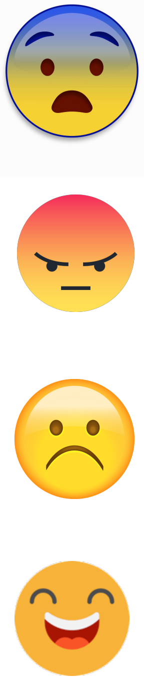
Hip Hop

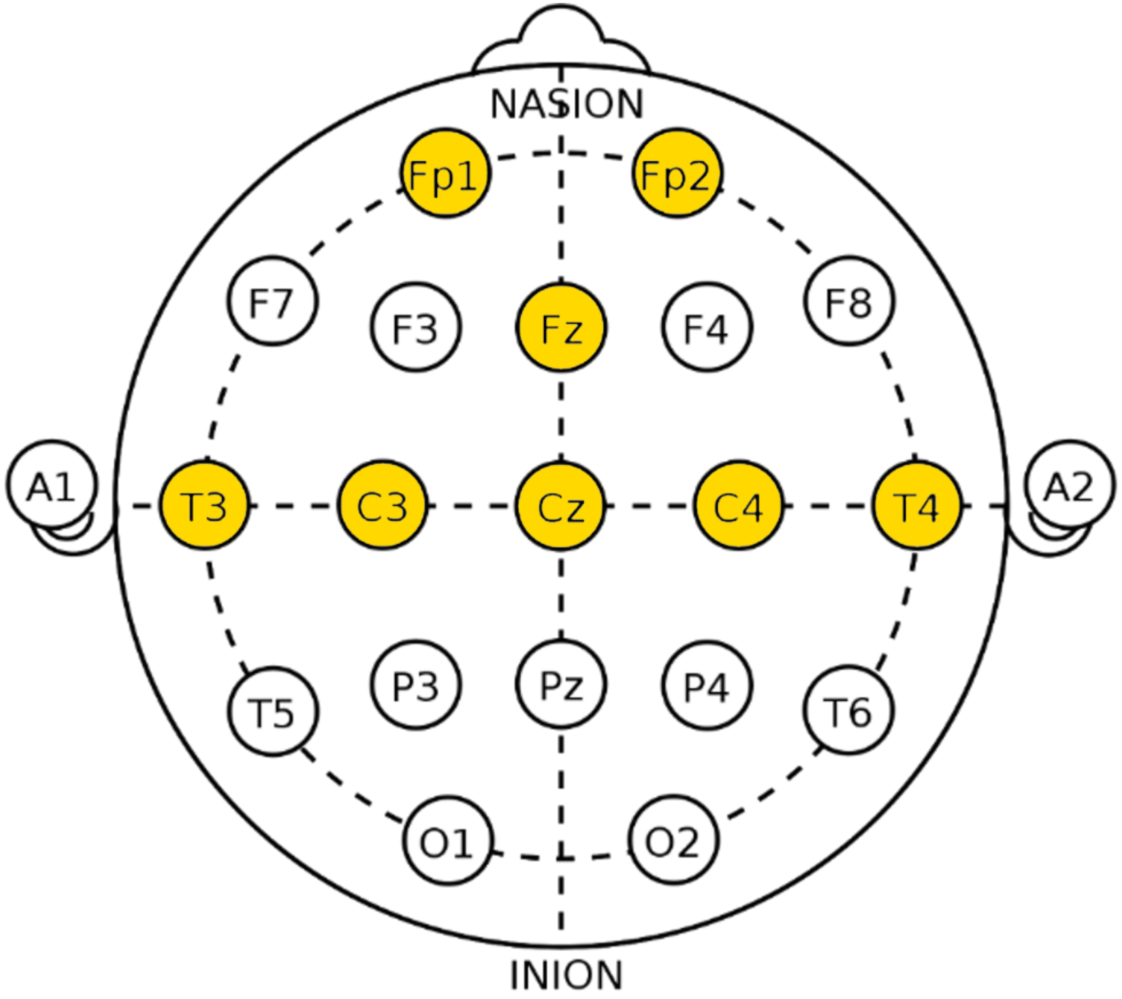
Classical

Popular

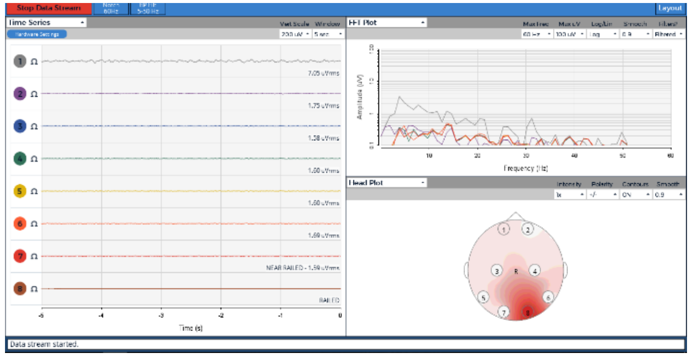
Heavy Metal

Electronic





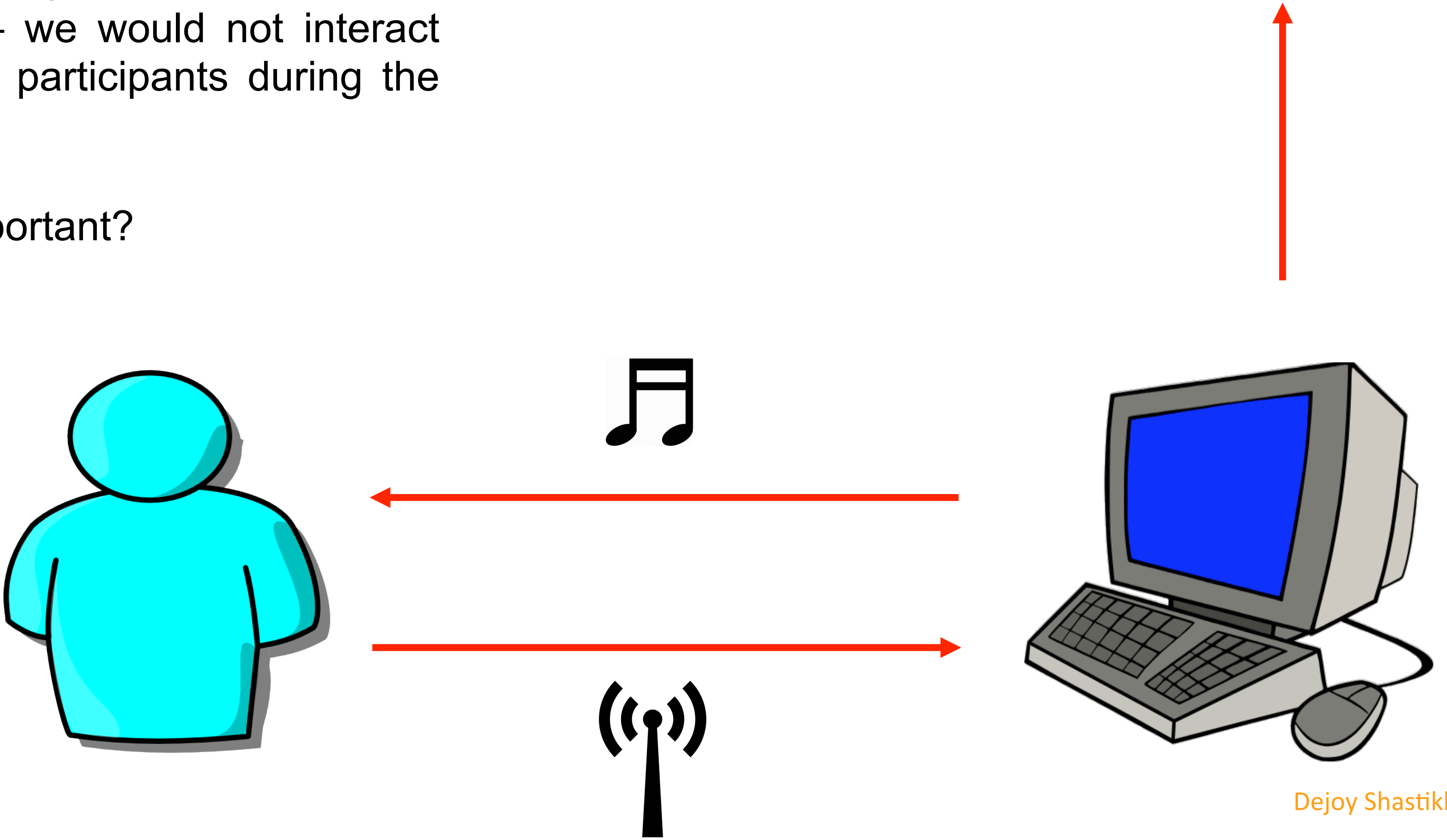
10-20 Electrode Placement Map

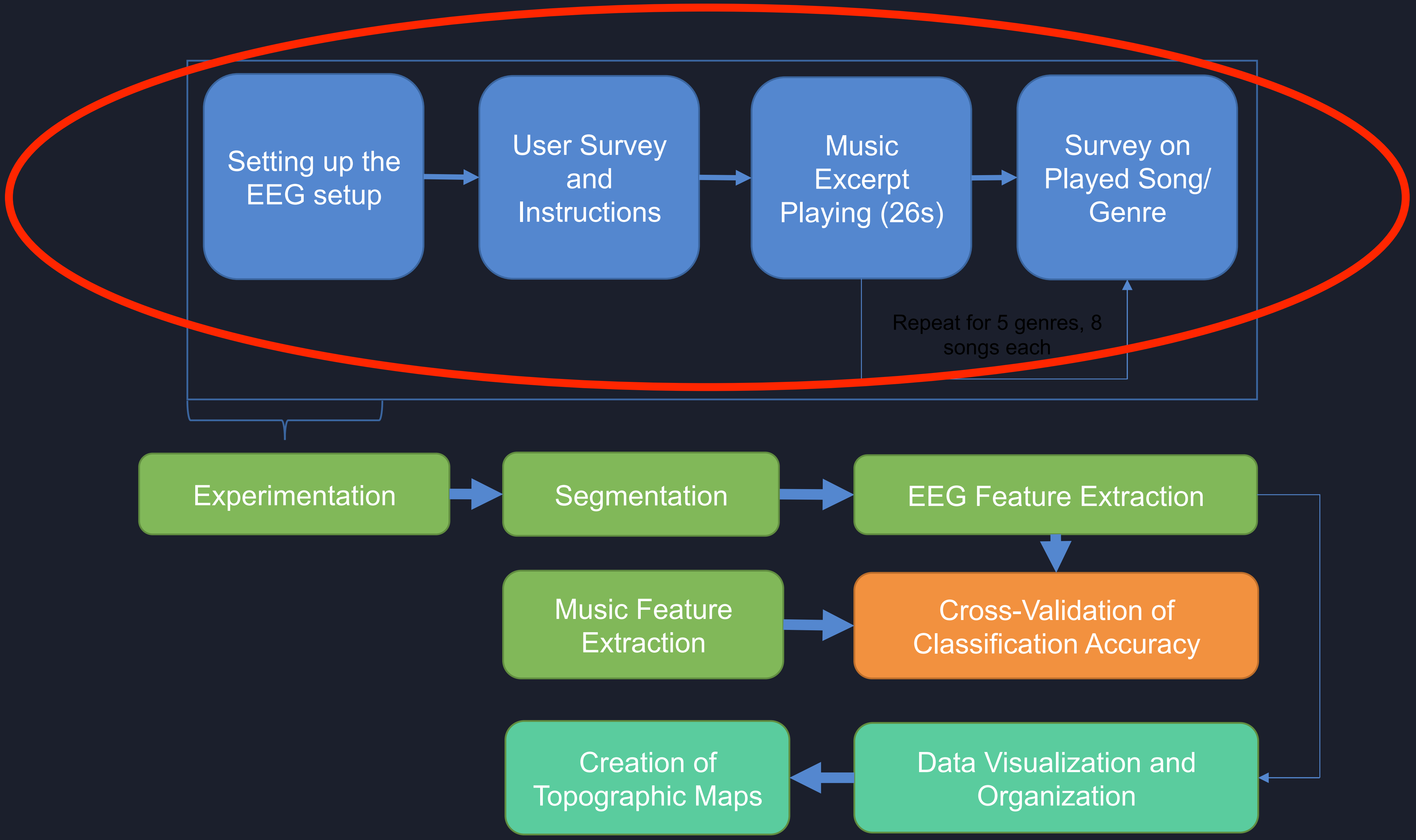


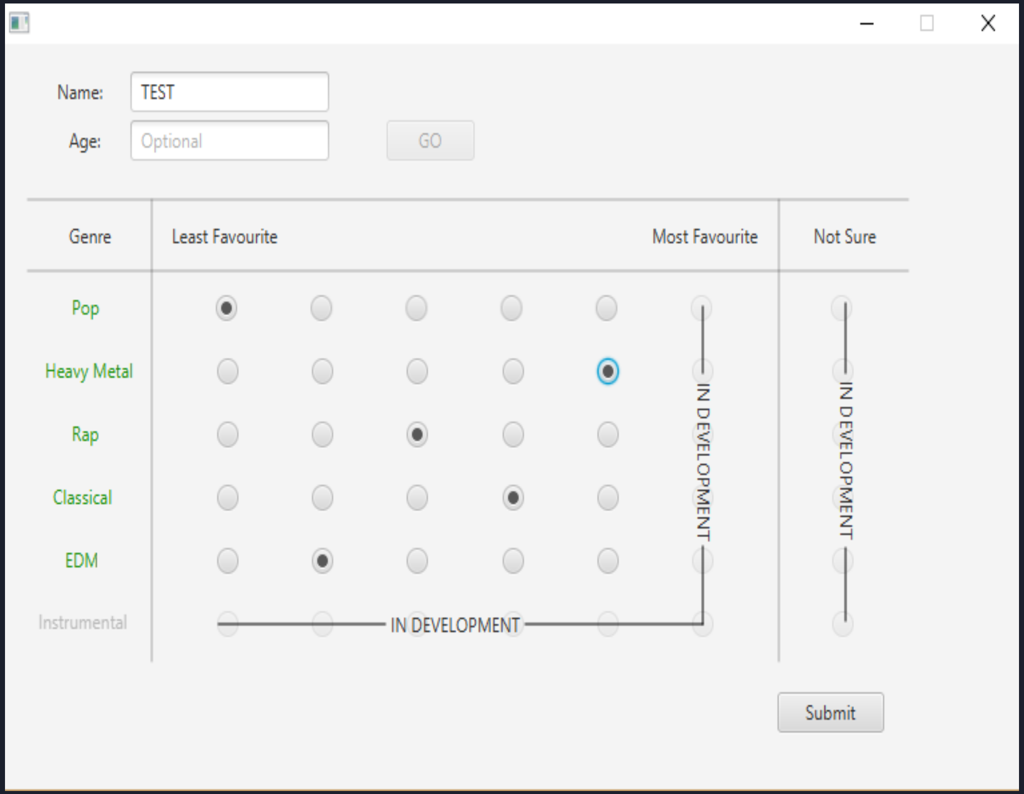
Experimental Program

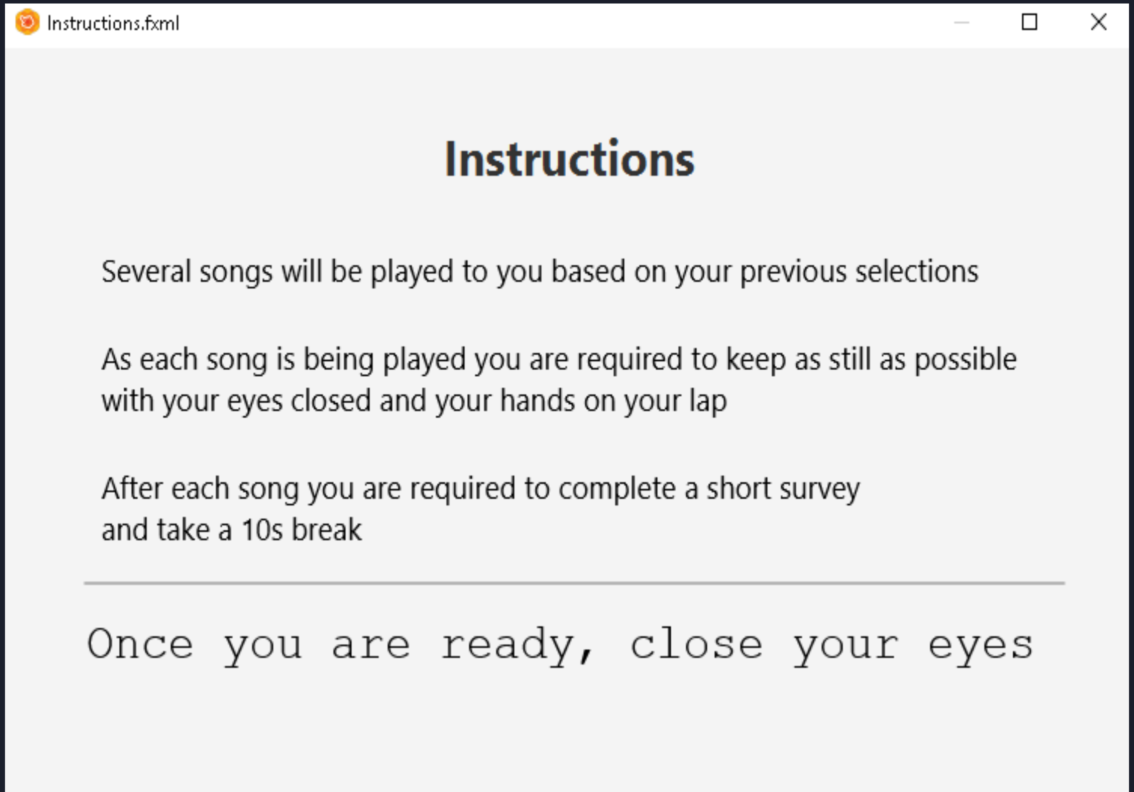
We introduced a **fully automated** experimental program for **experiment participants** – we would not interact **at all** with the participants during the experiment.

Why is this important?

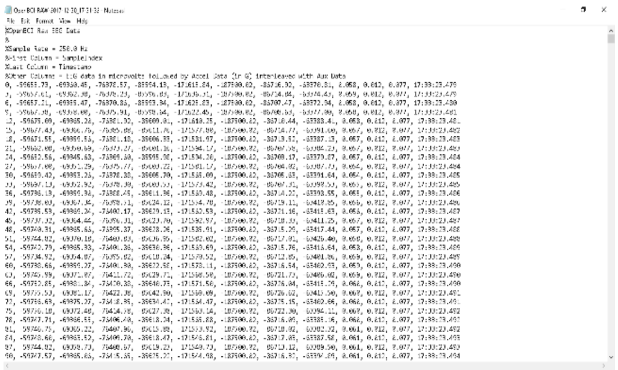



Setting up the EEG setup
User Survey and Instructions
Music Excerpt Playing (26s)
Survey on Played Song/ Genre
Repeat for 5 genres, 8 songs each
Experimentation
Segmentation
EEG Feature Extraction
Music Feature Extraction
Cross-Validation of Classification Accuracy
Creation of Topographic Maps
Data Visualization and Organization


Name:
TEST
Age:
Optional
GO
Genre
Least Favourite
Most Favourite
Not Sure
Pop
Heavy Metal
Rap
Classical
EDM
Instrumental
IN DEVELOPMENT
IN DEVELOPMENT
IN DEVELOPMENT
Submit

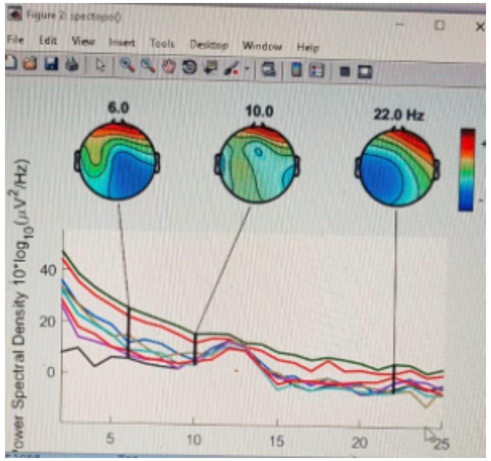

Instructions.fxml
Instructions
Several songs will be played to you based on your previous selections
As each song is being played you are required to keep as still as possible with your eyes closed and your hands on your lap
After each song you are required to complete a short survey and take a 10s break
Once you are ready, close your eyes

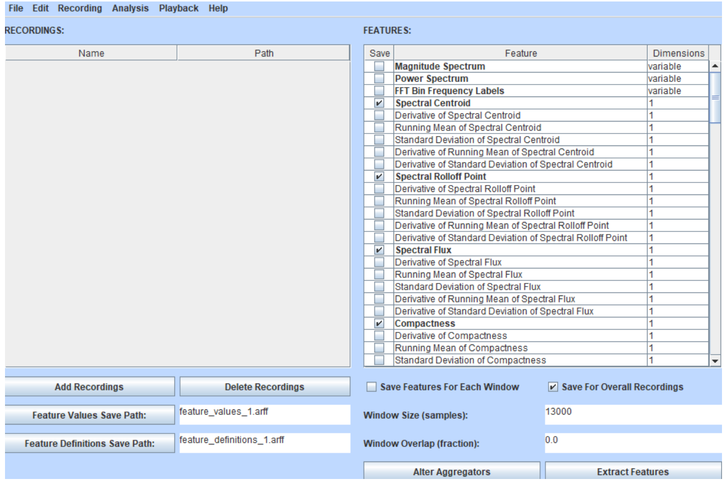
Segmentation and Extraction of Features



EEGrunt

Python utilities for EEG data analysis + display.





Only **low-level features** are used in *jAudio* to demonstrate higher generalizability, as some high level features are not easily extracted.

Feature selection is done to select important features.

3

VALIDATION

Machine Learning Algorithms Featured

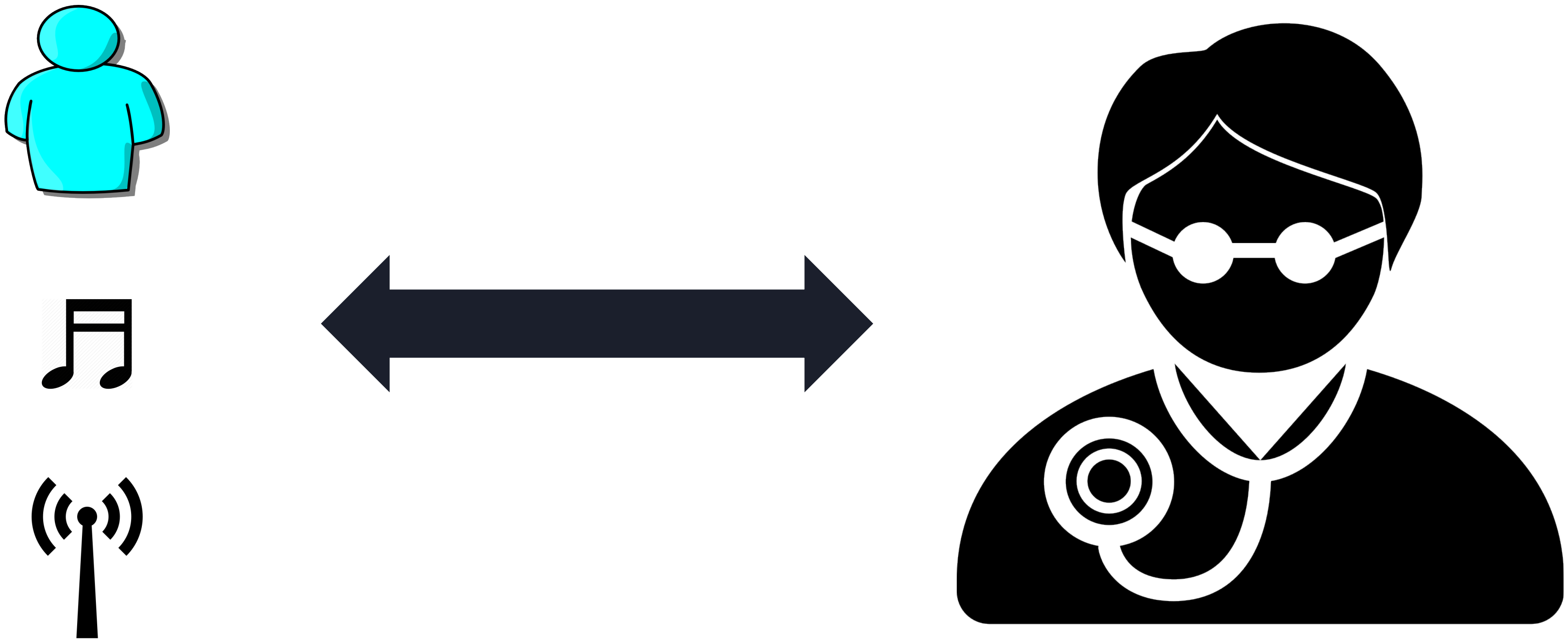
Random Forest

Can avoid overfitting, and can thus be highly accurate as it does not recognize noise. Averages multiple decision trees (this can make predictions harder to interpret)

Instance-based

Can adapt to unseen data – thus, highly useful for a forever expanding database of new data and new music. Potentially allows for individual user customization.

How We Validate Subjective Annotation and Machine Recognition of Music Emotion



Subjective Annotation,
Music Features, EEG Data

Expert Annotations on
Music Pieces

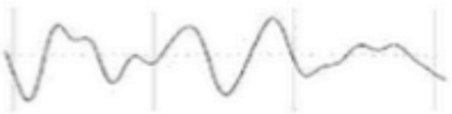
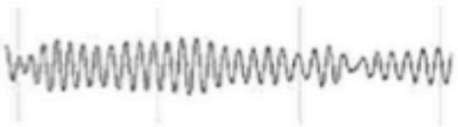
Frequency Bands

Useful for Emotion Analysis:

- **Alpha**
- **Low-Beta**
- **Beta**

Not Useful for Emotion Analysis:

- Delta (sleep analysis)
- Theta (semi-sleep states/ drowsiness)
- Gamma (sensory processing in the visual cortex)

Frequency Band Name	Frequency Bandwidth	State Associated with Bandwidth	Example of Filtered Bandwidth
Raw EEG	0–45 Hz	Awake	
Delta	0.5–3.5 Hz	Deep Sleep	
Theta	4–7.5 Hz	Drowsy	
Alpha	8–12 Hz	Relaxed	
Beta	13–35 Hz	Engaged	

10-fold Cross Validation Accuracies

Random Forest	Instance-based
98.2%	86.8%

Electroencephalography Data

Random Forest	Instance-based
95.0%	84.0%

Low-Level Music Features

Alpha (α: 8–13 Hz)	Low beta (β: 14–17 Hz)	Beta (β: 18–31 Hz)
90.4%	**91.2%**	88.0%

Most Predictive Frequency Bands

Why is the accuracy so high?

Instance-based classification allows for more precise calculation of emotion in each segment (less variation)

2/3 of all instances in a song have the same predicted music emotion → That song will be predicted to possess that said emotion

EEG Classification	Music Feature Classification	Subjective Annotation
100% (40)	100% (40)	74.2%

Accuracy of individual music classification methods (as compared to expert-annotated set)

If two-thirds of the instances within a song **agree on the predicted emotion**, that emotion will be listed as predicted for that song and compared to the expert-annotated data

For subjective annotation data, user responses are **normalized and classified** into emotions, after which they are compared to the expert-annotated data.

Integrated Model for Prediction of Music Emotion

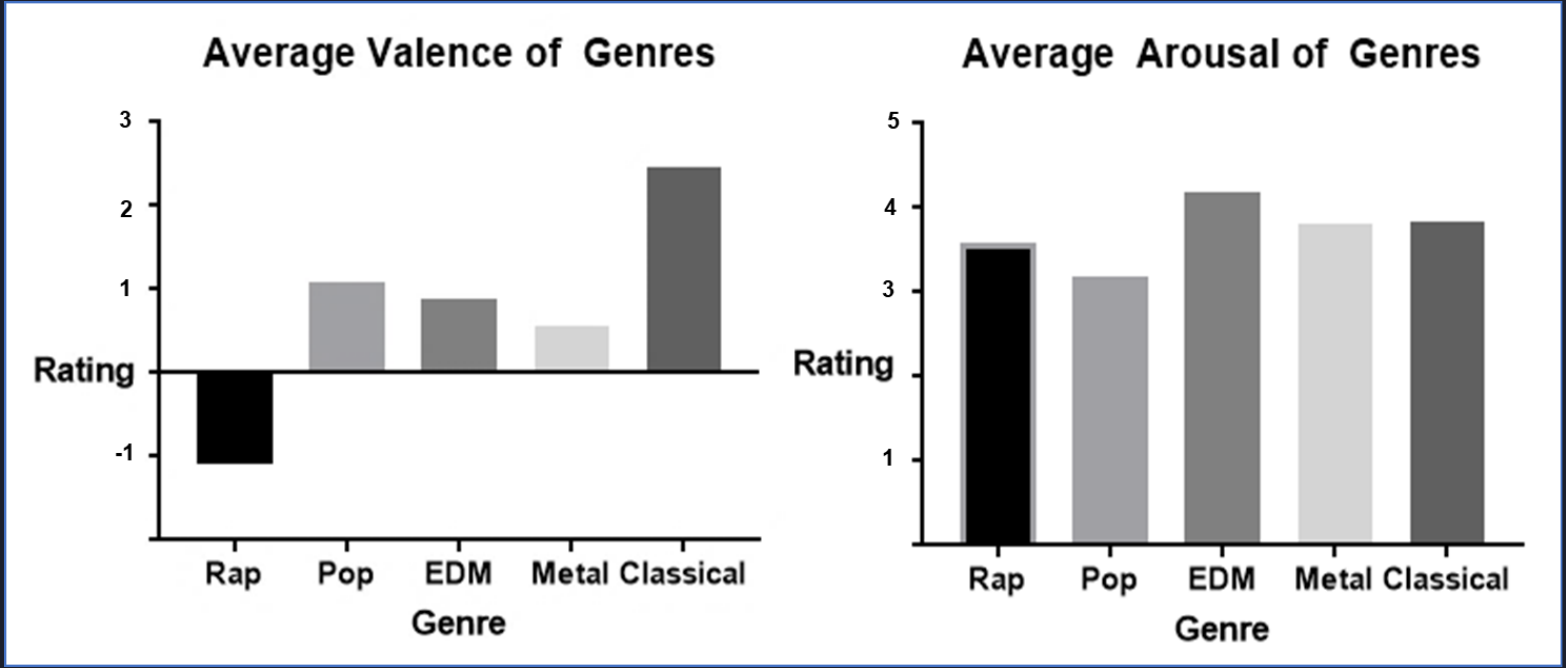
Prediction of emotion through low-level feature classifier

Use of expert-based annotation to annotate music (blinding used)

If both agree, song is classified as such. **2500 songs classified as such**

4

SECONDARY ANALYSIS



Genre preference **appears to be linked** to perceived positivity in music – which correlates to positive emotions as per the valence-arousal model

Type	Classical	Heavy Metal	Pop	Electronic	Rap
Preferred	70%	20%	60%	40%	10%
Disliked	10%	70%	20%	20%	80%

Preferred – Two most favorite genres among the five

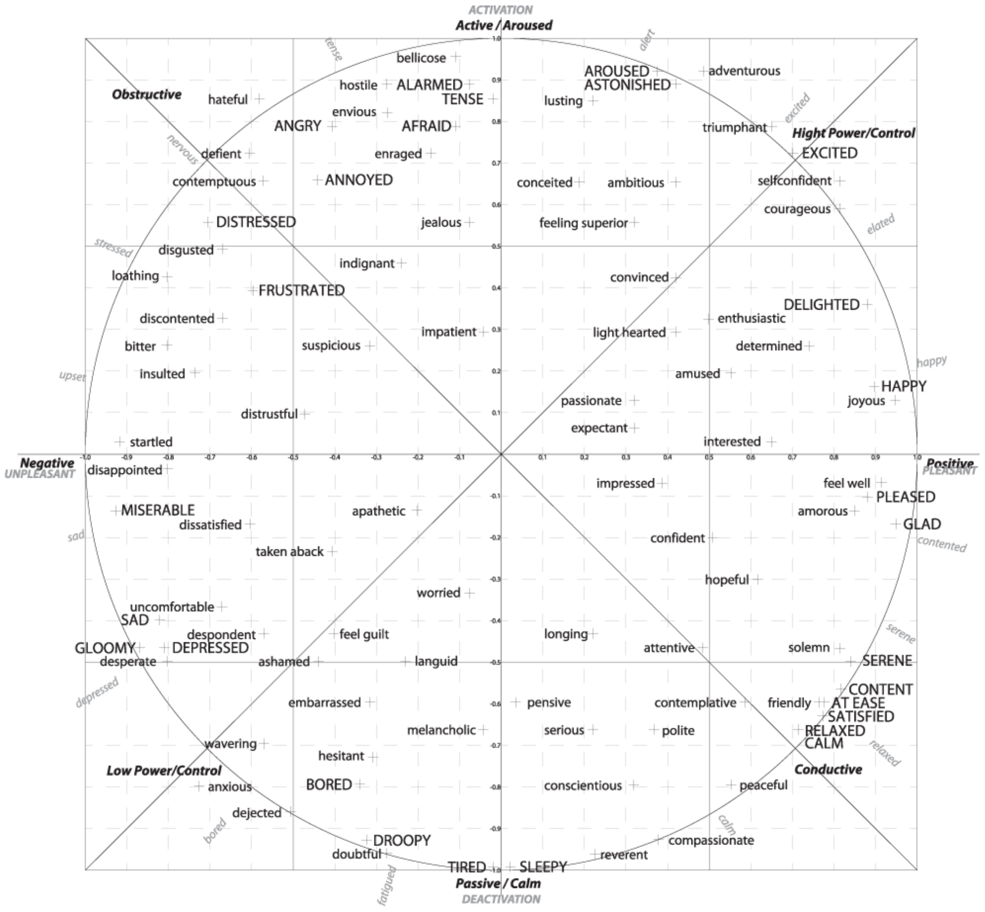
Disliked – Two least favorite genres among the five

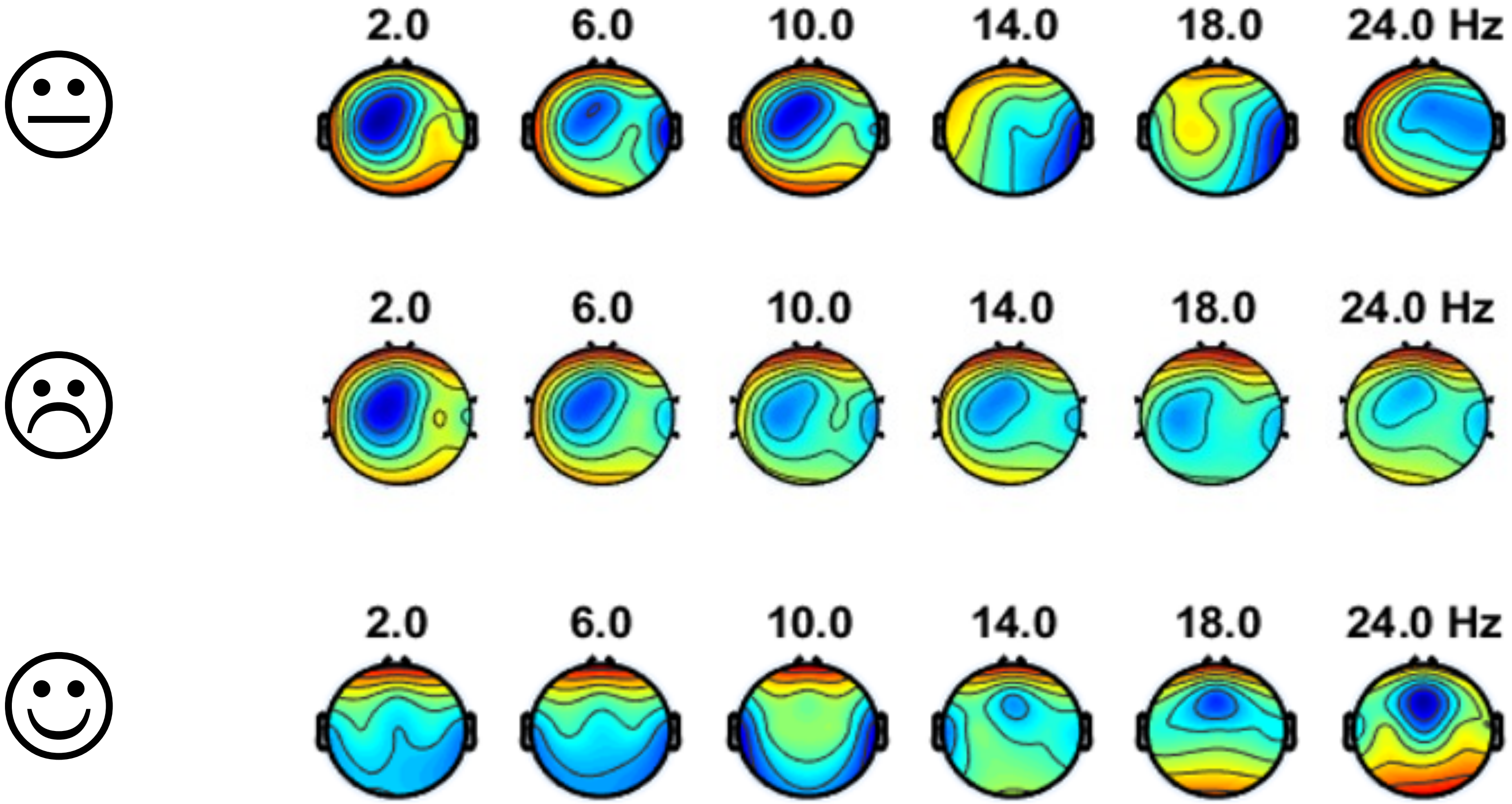
Is there a third axis to the emotion model?

We have shown from the **normalised average arousal and valence ratings** that for at least 2 of the genres there is indeed a **perceived link between valence and genre preference.**

However, we can **only propose that there is a third axis to the emotion model**, as it may itself be a confounding factor and we have **limited data to validate our work here as it was not our main objective.**

We suggest an **extension to this project** with a **diverse collection of songs per genre** to solidify that conclusion. The number of subjects also has to be increased to increase the reliability of results.





Valence
Alpha
Beta

Arousal
Intensity

Topographic Maps

MATLAB

EEGLAB
12th Workshop • November 2010

5

APPLICATION & VALIDATION

Integrated Model for Prediction of Music Emotion

Prediction of emotion through low-level feature classifier

Use of expert-based annotation to annotate music (blinding used)

If both agree, song is classified as such. **2500 songs classified as such**

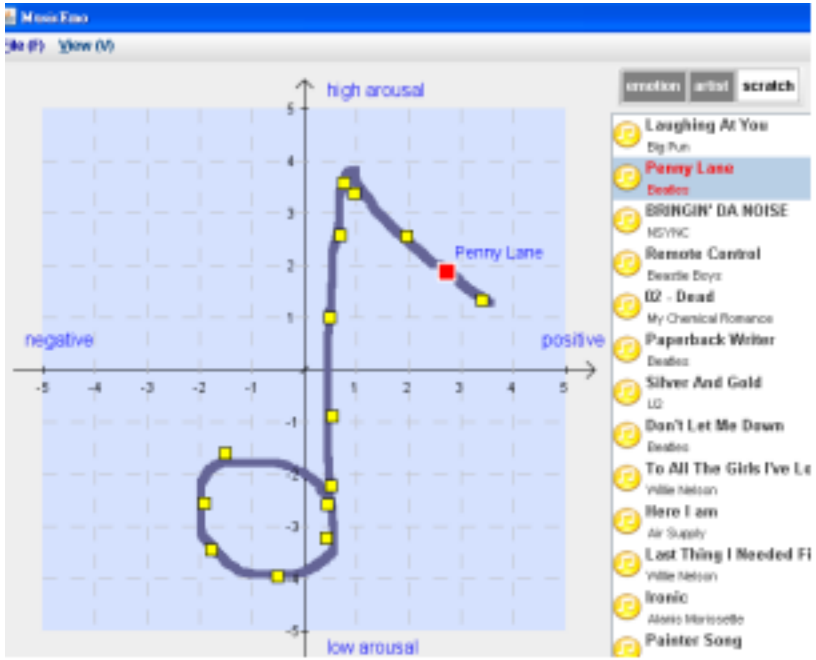
Bruno Mars
24K Magic

XXXTentacion
Moonlight

Problems with Existing Work

- Methodologies exist to go from music at **one end of the valence-arousal scale to another**
- Current work does not consider **psychological findings** to create **automatic playlists** for users to use
- **High complexity** on user's end – this is undesirable!
- Users may not understand how to **effect meaningful emotional change**, and this leads to worse results

(Mr Emo., National Taiwan University)

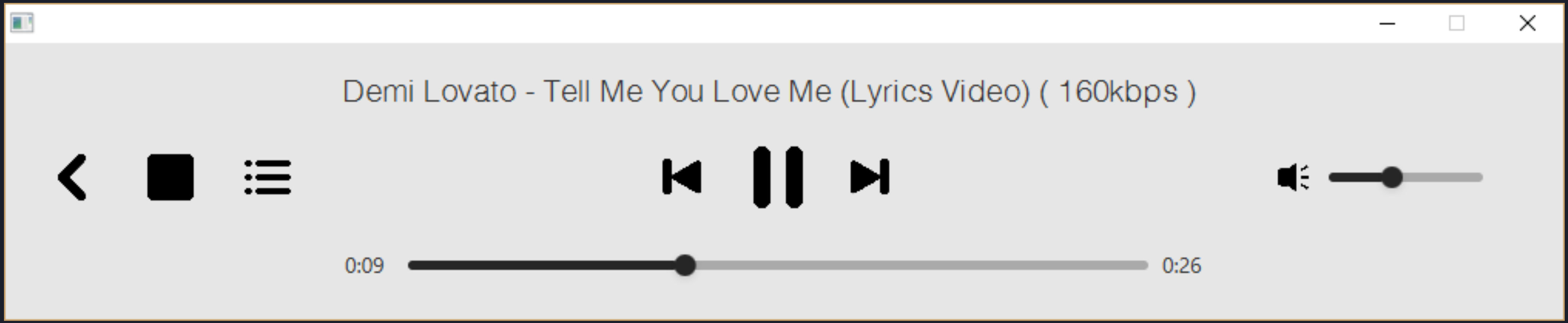


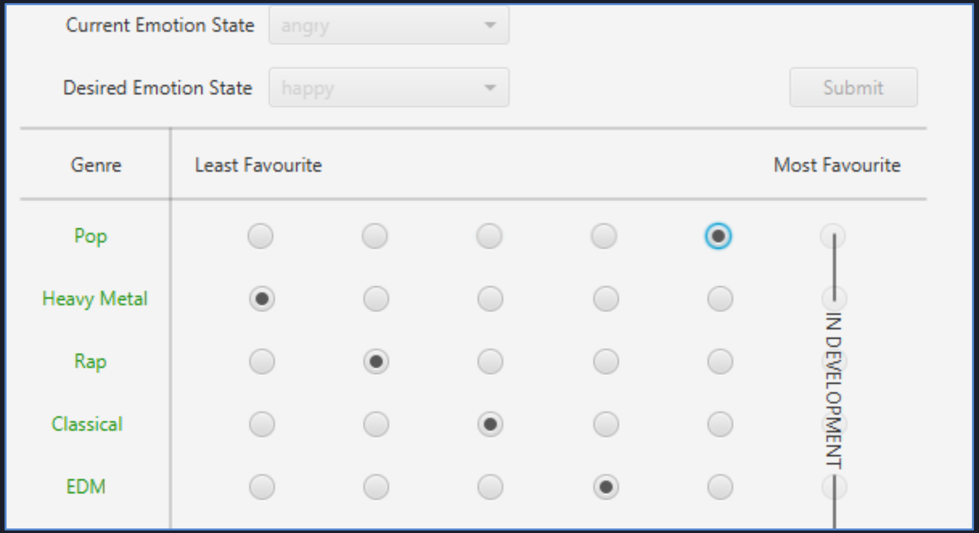
Music must **transition** from original to final emotion to **keep the user engaged**.

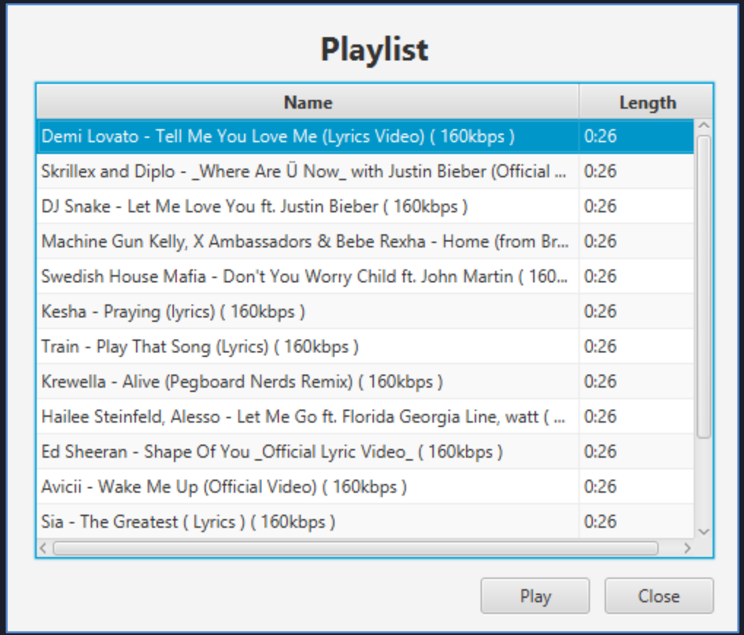
Lack of application for general public to **induce emotional change** for **therapeutic benefits**

Develop application to create musical playlists which will **induce desired mood** in users

Music must be **preferred** by listener to **emotionally engage** the listener


Demi Lovato - Tell Me You Love Me (Lyrics Video) (160kbps)
0:09
0:26


Current Emotion State
angry
Desired Emotion State
happy
Submit
Genre
Least Favourite
Most Favourite
Pop
Heavy Metal
Rap
Classical
EDM
IN DEVELOPMENT


Playlist
Name
Length
Demi Lovato - Tell Me You Love Me (Lyrics Video) (160kbps)
0:26
Skrillex and Diplo - _Where Are Ü Now_ with Justin Bieber (Official ...
0:26
DJ Snake - Let Me Love You ft. Justin Bieber (160kbps)
0:26
Machine Gun Kelly, X Ambassadors & Bebe Rexha - Home (from Br...
0:26
Swedish House Mafia - Don't You Worry Child ft. John Martin (160...
0:26
Kesha - Praying (lyrics) (160kbps)
0:26
Train - Play That Song (Lyrics) (160kbps)
0:26
Krewella - Alive (Pegboard Nerds Remix) (160kbps)
0:26
Hailee Steinfeld, Alesso - Let Me Go ft. Florida Georgia Line, watt (...
0:26
Ed Sheeran - Shape Of You _Official Lyric Video_ (160kbps)
0:26
Avicii - Wake Me Up (Official Video) (160kbps)
0:26
Sia - The Greatest (Lyrics) (160kbps)
0:26
Play
Close

Validation of our Work

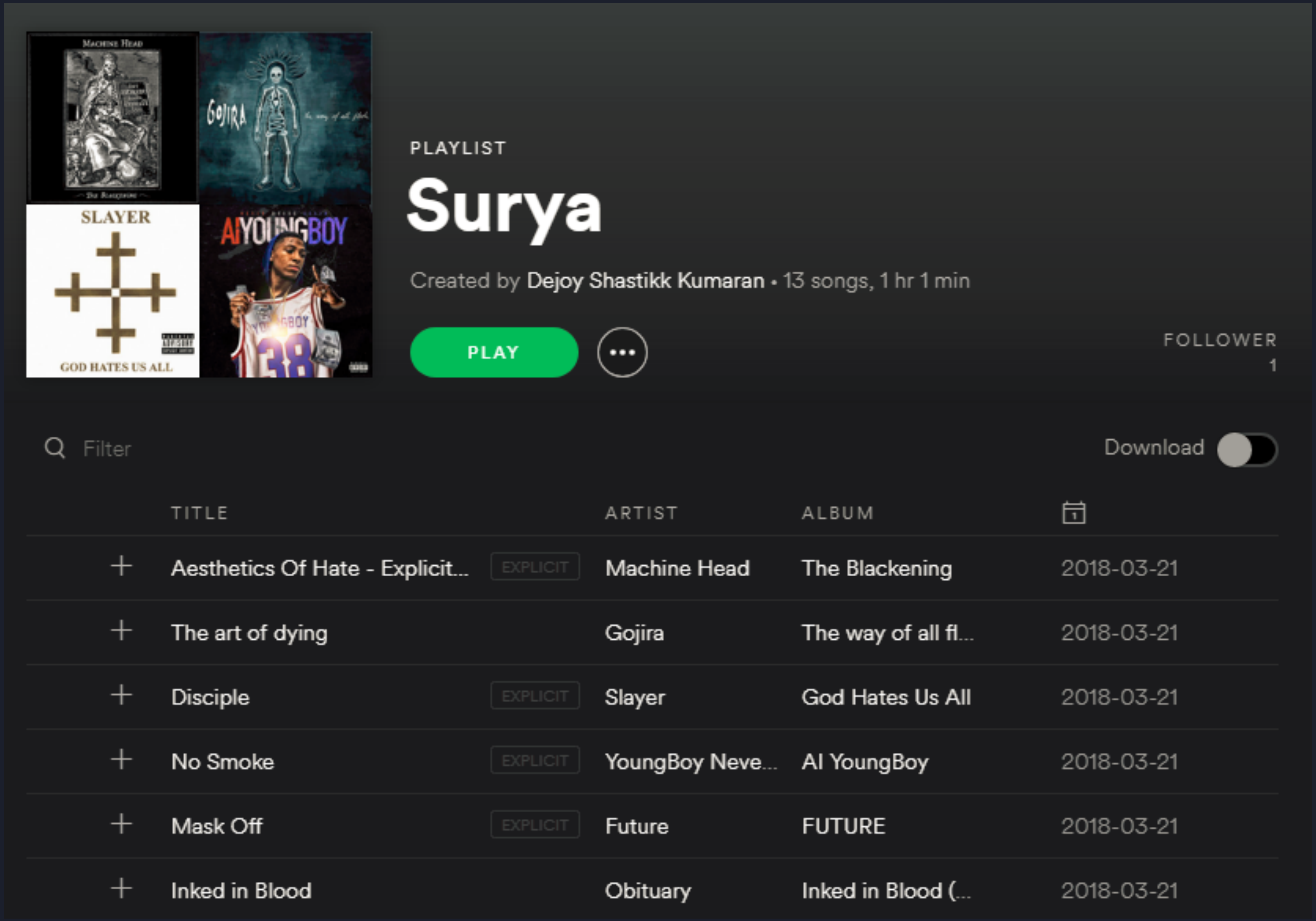
There were 3 groups for the experiment – a control group, a group which used only 1 genre and a group which used 2 preferred genres.

There were 31 participants, with an age range of 36 and a median age of 18. Participants did not have more than a year of music education.

Custom playlists were created **on popular streaming platforms**

A total of 31 users were asked to listen to the playlist **after providing** a mood rating (upon 10)

After listening to the playlist, they were asked to **rate their mood** again


PLAYLIST
Surya
Created by Dejoy Shastikk Kumaran • 13 songs, 1 hr 1 min
PLAY
FOLLOWER
1
Filter
Download
TITLE
ARTIST
ALBUM
Aesthetics Of Hate - Explicit...
EXPLICIT
Machine Head
The Blackening
2018-03-21
The art of dying
Gojira
The way of all fl...
2018-03-21
Disciple
EXPLICIT
Slayer
God Hates Us All
2018-03-21
No Smoke
EXPLICIT
YoungBoy Neve...
AI YoungBoy
2018-03-21
Mask Off
EXPLICIT
Future
FUTURE
2018-03-21
Inked in Blood
Obituary
Inked in Blood (...
2018-03-21

Results of Validation

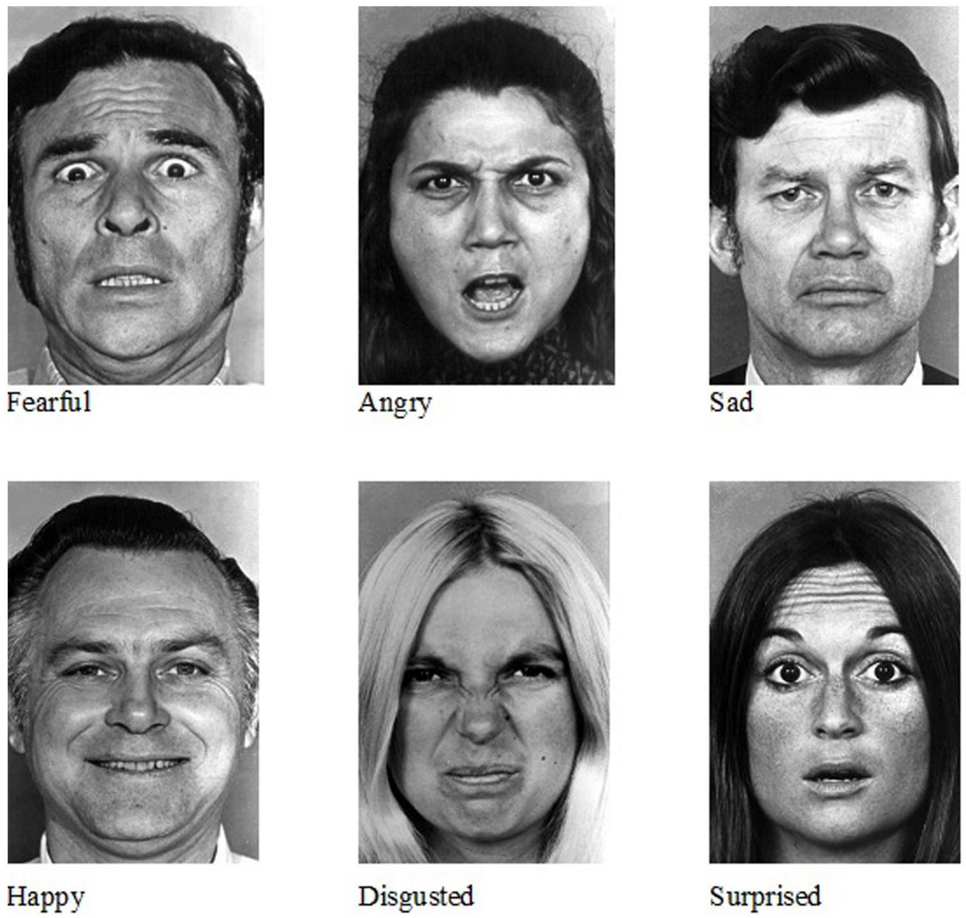
	Happy Songs Only	Custom (2 Genres)	Custom (1 Genre)
Mean Change	1.13	2.58	**4.18**
Standard Deviation	1.64	1.16	1.25
P-Value	-	**0.016**	**0.0001**

Validated that our custom methodology provides **more positive mood change** for users (1 genre preferred)

5

CONCLUSION, FUTURE WORK AND REFERENCES

- Obtained **very high** classification accuracies with EEG and Music Feature Data
- Validated **subjective annotation** and **low-level features** for machine recognition uses
- Analysed links between **genre preference** and **perceived valence and arousal**
- Created **topographic maps** to analyse links between **frequency bands and emotion**
- Designed **music playlist creation technique** to induce desired emotion
- Classified **2500 songs** using an **integrated approach** in a validated application
- **Commercialization of technology** alongside **extension to more genres** and affects
- Development of **single-image emotion recognition technology** to remove need for manual emotional input



References

1. Aha, D. W., Kibler, D., & Albert, M. K. (1991). Instance-based learning algorithms. Machine learning, 6(1), 37-66.
2. Bailey, L. M. (1984). The use of songs in music therapy with cancer patients and their families. Music Therapy, 4(1), 5-17.
3. Barthet, M., Fazekas, G., & Sandler, M. (2012, June). Music emotion recognition: From content-to context-based models. In International Symposium on Computer Music Modeling and Retrieval (pp. 228-252). Springer, Berlin, Heidelberg.
4. Bruner, G. C. (1990). Music, mood, and marketing. the Journal of marketing, 94-104.
5. EEGrunt. (2017, June 19). Retrieved January 06, 2018, from https://github.com/curiositry/EEGrunt
6. Fornäs, J. (2006). Media passages in urban spaces of consumption.

1. G.E. Chatrian, E. Lettich, and P.L. Nelson. Ten percent electrode system for topographic studies of spontaneous and evoked EEG activity. Am J EEG Technol, 25:83-92, 1985.
2. Gevins, A., Smith, M. E., McEvoy, L., & Yu, D. (1997). High-resolution EEG mapping of cortical activation related to working memory: effects of task difficulty, type of processing, and practice. Cerebral cortex (New York, NY: 1991), 7(4), 374-385.
3. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. ACM SIGKDD explorations newsletter, 11(1), 10-18.
4. Haq, S., Jackson, P. J., & Edge, J. (2008). Audio-visual feature selection and reduction for emotion classification. In Proc. Int. Conf. on Auditory-Visual Speech Processing (AVSP’08), Tangalooma, Australia.

References

1. Laurier, C., Grivolla, J., & Herrera, P. (2008, December). Multimodal music mood classification using audio and lyrics. In Machine Learning and Applications, 2008. ICMLA'08. Seventh International Conference on (pp. 688-693). IEEE.
2. Li, T., Ogihara, M., & Li, Q. (2003, July). A comparative study on content-based music genre classification. In Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval (pp. 282-289). ACM.
3. Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T. L., Jeng, S. K., Duann, J. R., & Chen, J. H. (2010). EEG-based emotion recognition in music listening. IEEE Transactions on Biomedical Engineering, 57(7), 1798-1806.
4. Lin, Y. P., & Jung, T. P. (2017). Improving EEG-Based Emotion Classification Using Conditional Transfer Learning. Frontiers in human neuroscience, 11, 334.

1. Lu, L., Liu, D., & Zhang, H. J. (2006). Automatic mood detection and tracking of music audio signals. IEEE Transactions on audio, speech, and language processing, 14(1), 5-18.
2. MacDorman, Stuart Ough Chin-Chang Ho, K. F. (2007). Automatic emotion prediction of song excerpts: Index construction, algorithm design, and empirical comparison. Journal of New Music Research, 36(4), 281-299.
3. Magee, W. L., & Davidson, J. W. (2002). The effect of music therapy on mood states in neurological patients: a pilot study. Journal of Music Therapy, 39(1), 20-29.
4. McEnnis, D., C. McKay, and I. Fujinaga. 2006. jAudio: Additions and improvements. Proceedings of the International Conference on Music Information Retrieval. 385–6.

References

1. McEnnis, D., C. McKay, and I. Fujinaga. 2006. Overview of OMEN. Proceedings of the International Conference on Music Information Retrieval. 7–12.
2. McEnnis, D., C. McKay, I. Fujinaga, and P. Depalle. 2005. jAudio: A feature extraction library. Proceedings of the International Conference on Music Information Retrieval. 600–3.
3. McKay, C. 2010. Automatic music classification with jMIR. Ph.D. Thesis. McGill University, Canada.
4. McKay, C., J. A. Burgoyne, J. Hockman, J. B. L. Smith, G. Vigliensoni, and I. Fujinaga. 2010. Evaluating the genre classification performance of lyrical features relative to audio, symbolic and cultural features. Proceedings of the International Society for Music Information Retrieval Conference. 213–8.
5. McKay, C., and I. Fujinaga. 2010. Improving automatic music classification performance by extracting features from different types of data. Proceedings of the ACM SIGMM International Conference on Multimedia Information Retrieval. 257–66.

1. McKay, C., and I. Fujinaga. 2010. Improving automatic music classification performance by extracting features from different types of data. Proceedings of the ACM SIGMM International Conference on Multimedia Information Retrieval. 257–66.
2. McKay, C., and I. Fujinaga. 2009. jMIR: Tools for automatic music classification. Proceedings of the International Computer Music Conference. 65–8.
3. McKay, C., J. A. Burgoyne, and I. Fujinaga. 2009. jMIR and ACE XML: Tools for performing and sharing research in automatic music classification. Presented at the ACM/IEEE Joint Conference on Digital Libraries Workshop on Integrating Digital Library Content with Computational Tools and Services.
4. McKay, C., and I. Fujinaga. 2008. Combining features extracted from audio, symbolic and cultural sources. Proceedings of the International Conference on Music Information Retrieval. 597–602.

References

1. Prinzie, A., Van den Poel, D. (2008). "Random Forests for multiclass classification: Random MultiNomial Logit". Expert Systems with Applications. 34 (3):
2. Prism, G. (1994). Graphpad software. San Diego, CA, USA.
3. Ruck, D. W., Rogers, S. K., Kabrisky, M., Oxley, M. E., & Suter, B. W. (1990). The multilayer perceptron as an approximation to a Bayes optimal discriminant function. IEEE Transactions on Neural Networks, 1(4), 296-298.
4. Russell, James (1980). "A circumplex model of affect". Journal of Personality and Social Psychology. 39: 1161–1178.
5. Särkämö, T., Tervaniemi, M., Laitinen, S., Forsblom, A., Soinila, S., Mikkonen, M., ... & Peretz, I. (2008). Music listening enhances cognitive recovery and mood after middle cerebral artery stroke. Brain, 131(3), 866-876.

1. Thornton, C., Hutter, F., Hoos, H. H., & Leyton-Brown, K. (2013, August). Auto-WEKA: Combined selection and hyperparameter optimization of classification algorithms. In Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 847-855). ACM.
2. Tolos, M., Tato, R., & Kemp, T. (2005, January). Mood-based navigation through large collections of musical data. In Consumer Communications and Networking Conference, 2005. CCNC. 2005 Second IEEE (pp. 71-75). IEEE.
3. Trohidis, K., Tsoumakas, G., Kalliris, G., & Vlahavas, I. P. (2008, September). Multi-Label Classification of Music into Emotions. In ISMIR (Vol. 8, pp. 325-330).

References

1. Wieczorkowska, A., Synak, P., Lewis, R., & Raś, Z. W. (2005, May). Extracting emotions from music data. In International Symposium on Methodologies for Intelligent Systems (pp. 456-465). Springer, Berlin, Heidelberg.
2. Wu, T. L., & Jeng, S. K. (2006). Automatic emotion classification of musical segments. In Proceedings of the 9th International Conference on Music Perception & Cognition, Bologna.
3. Yang, Y.-H. and Chen, H. H. 2012. Machine recognition of music emotion: A review. ACM Trans. Intell. Syst. Technol. 3, 3, Article 40 (May 2012), 30 pages
4. Yang, Yi-Hsuan, Chia-Chu Liu, and Homer H. Chen. "Music emotion classification: A fuzzy approach." In Proceedings of the 14th ACM international conference on Multimedia, pp. 81-84. ACM, 2006.

1. Yang, Y. H., Lin, Y. C., Cheng, H. T., & Chen, H. H. (2008, October). Mr. Emo: Music retrieval in the emotion plane. In Proceedings of the 16th ACM international conference on Multimedia (pp. 1003-1004). ACM.
2. Lin, Y. P., Wang, C. H., Wu, T. L., Jeng, S. K., & Chen, J. H. (2009, April). EEG-based emotion recognition in music listening: A comparison of schemes for multiclass support vector machine. In Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on (pp. 489-492). IEEE.

Thanks!

Connect with me at
shastikk@gmail.com or **shastikk@ieee.org**

5

FURTHER QUESTIONS

Why did you pick the machine learning algorithms selected?

- Existing papers often **use Support Vector Machines (SVMs)** so we wanted to explore a **wider range** of classification algorithms

- **High accuracy** achieved for these classifiers in similar prediction applications and research

- The **advantages of each machine learning algorithm** as earlier stated – instance-based **are able to adapt to unseen data** and random forests **avoid overfitting**

What is the research standard for classification accuracy in your experiment?

- We have achieved very strong results, with accuracy of **>98% for EEG data** as compared to the research standard of **~94% with SVM.**

- We have also achieved strong results for **music feature classification at the low level,** achieving **95% in comparison to the 66-70% research standard,** perhaps due to the splitting of each song into instances

Why does the initial emotion state's songs need to be played if we are intending the user to move to the final emotion?

- As Bailey states, alongside multiple similar research papers, users are **more emotionally engaged** when the music they are listening to is **emotionally similar to them.**

- Important to **engage users** in order to get them to listen to the music for longer with **higher effectiveness in engagement.**